Standardized Nursing Classification Systems:
Necessary, but Not Sufficient, for Representing What Nurses Do

Suzanne Bakken Henry, RN, DNSc[1], Charles N. Mead, MD, MS[1,2]
[1]Department of Community Health Systems, University of California, San Francisco, CA
[2]CareCentric Solutions, Inc., Atlanta, Georgia

The American Nurses Association Steering Committee on Databases to Support Nursing Practice has recognized three standardized nursing intervention classification systems. Because each of these classifications systems focuses on encoding informational abstractions of nursing actions, rather than providing a controlled vocabulary and compositional grammar from which informational concepts and abstractions can be constructed, the systems are necessary, but not sufficient, for representing what nurses do. In particular, computer-based patient record systems focused on process understanding and process improvement will require atomic-level representations of nursing actions suitable for transformation into a variety of information abstractions, including, but not limited to, the abstractions contained in the three existing classification systems.

INTRODUCTION

The healthcare delivery system in the United States is being reformed, reshaped and reorganized. Hardly a day goes by in which one does not hear about new strategic alliances, mergers, acquisitions, and re-definitions of the roles and responsibilities of providers, payors and patients. As the healthcare delivery system marches inexorably down the road to "managed care," the underlying assumptions that have traditionally motivated care delivery are being challenged with requests to justify the appropriateness and effectiveness of the care. At the core of these challenges is a mandated administrative requirement to control the cost of healthcare delivery. This "bottom line" perspective is offset by a seemingly contradictory clinical requirement to maintain (and hopefully improve) the quality of care, and additionally to make that care available to all who need it.

"Healthcare information systems" have traditionally focused on capturing and analyzing data about the utilization of the delivery system – the administrative processes of patient care. Information systems to capture and analyze information about what clinicians *do* (as opposed to what they bill for), have lagged significantly behind administrative systems. This is due in part to the historically strong emphasis that healthcare information systems have placed in "charge capture," and in part to the inherent complexity of clinical processes.

As a result of the presence of information systems containing vast amounts of administrative process data and the absence of systems containing clinical process data, what we call "managed care" in 1996 can more realistically be called "managed utilization." The advent of truly "managed" patient care awaits the arrival of systems capable of representing the details of clinical processes in a form that will allow the data to be retrospectively analyzed, prospectively utilized, and meaningfully integrated into administrative decision making.

In this paper, we argue that, from the perspective of information content, effective capture of the details of the clinical actions taken by nurses will require not only the existing classification systems, but also taxonomic vocabularies capable of representing atomic-level nursing actions. We believe that such vocabularies are essential to the development of a multi-purpose unified nursing language system.

In this discussion, we first provide a critical analysis of three standardized systems for classifying nursing interventions using both the Ingenerf typology of taxonomic vocabularies for healthcare,[1] and Cimino's nine criteria for a multi-purpose controlled vocabulary for clinical information systems[2] as metrics for understanding the strengths, as well as the weaknesses, of the existing nursing classification systems. Secondly, we illustrate the loss of potentially significant atomic data through the use of each of the classification systems as a means of illustrating the point that the existing classification systems are *information*, rather than *data*, encoding methodologies. Finally, we recommend areas for future research and development.

CRITICAL ANALYSIS

The American Nurses Association Steering

Committee on Databases to Support Nursing Practice has established several criteria that a classification system must satisfy to be recognized by the Committee.[3] The criteria focus upon the clinical usefulness, reliability and validity, and processes for revision and extension of the classification system. Additionally, the criteria state that all terms used by the system must be "defined precisely," and that the system must provide a unique identifier or code for each term.

To date, three intervention systems have been recognized through this program: 1) Omaha Community Health System[4] 2) Georgetown Home Health Care Classification[5]; and 3) Nursing Intervention Classification[6]. Several other systems are currently under evaluation.[7-8] For the purposes of this discussion, we present examples using the three recognized systems.

Typology of Taxonomic Vocabularies
Ingenerf has explicated four types of taxonomic vocabularies for healthcare based on the underlying structure and related knowledge representation formalism. **Thesauri** are defined as lexical vocabularies containing definitions and cross-references (e.g. the UMLS Metathesaurus).
Classification Systems are vocabularies which can be represented as hierarchies or decision trees, and which have as a main emphasis the disjunctive and exhaustive classification of terms. The International Classification of Diseases, and the Diagnostic and Statistical Manual are examples of Classification Systems. **Nomenclatures** are combinatorial taxonomic vocabularies containing more complex polyhierarchies or axes. Terms within a nomenclature may be combined into complex concepts using semantic grammars; however, explicit rules for canonical representation of terms are lacking. SNOMED International is considered to be an example of a Nomenclature. Finally, **Formal Terminologies**, such as the GRAIL representation language developed in conjunction with the GALEN Project[9], are vocabularies which are based on concepts, rather than on terms and which include explicit rules for sensible composition of primitive concepts into complex concepts. The concepts are represented using knowledge formalisms such as description logic or conceptual graphs.

The Omaha Community Health System can be placed somewhere between a Classification System a and Nomenclature. As a Classification System, Omaha contains the requisite hierarchical organization. Furthermore, the inclusion of the intervention "Other" guarantees, at least from a user's perspective, the fact that the system is exhaustive. Relative to its classification as a Nomenclature, Omaha can be considered to be biaxial in that it includes an axis of "target" of intervention, as well as a modifier axis for use in describing the method of delivery of the target intervention. The modifier axis is referred to as a "Category" in the Omaha system. Categories include health teaching, guidance, and counseling; treatments and procedures; case management; and surveillance. However, although it is biaxial, the system does not contain a semantic grammar which would allow the combination of terms to express complex semantic concepts. Omaha therefore lacks the full expressive power of a combinatorial taxonomic vocabulary, and hence does not fully qualify as an Nomenclature.

The Georgetown Home Health Care Classification is conceptually and architecturally similar to the Omaha System, and thus fits somewhere between a Classification System and a Nomenclature in the Ingenerf typology. Similar to Omaha, it has a biaxial structure which allow nursing interventions to have associated "modifiers" of assess, teach, provide direct care, and manage. Like Omaha, Georgetown does not qualify as a combinatorial vocabulary (i.e. a Nomenclature) because it does not have an associated semantic grammar.

As its name and design purpose imply, the Nursing Interventions Classification (NIC) can be most correctly be viewed as a Classification System. Although NIC is more explicit is defining the populations of its respective informational clusters, it, like Omaha and Georgetown, lacks a semantic grammar, and ultimately has the clustering (abstraction) of nursing actions as its primary *raison d'etre*.

In summary, none of the three recognized intervention classification systems meet the Ingenerf definition of a Formal Terminology, thereby suggesting the need to augment the existing nursing intervention classification systems with more finely granular and formally-based taxonomic vocabularies and their associated knowledge representations. Such vocabularies would enable the type of data abstraction manifest in the existing classification systems to occur without the data loss that results from the lack of an underlying formalism. The following three

exemplars of data loss using the existing classification systems underscore the need for such a formalism.

Exemplars of Data Loss
Omaha Community Health System
The Omaha System consists of 62 "target" interventions (63 if one includes the target "Other") and, as mentioned above, each target can be logically modified through its placement in one of four "categories" which describe the target intervention's mode of nursing action.

Nursing interventions are viewed as "descriptive bins" which abstract the actions which nurses actually perform. Thus, Omaha contains targets such as "01.Anatomy and Physiology – structure and function of the human body," "07.Cardiac Care – activities directed toward maintenance of cardiac or circulatory function, including diet, medication, vital signs, and relief of edema," and "51.Skin care – activities directed toward maintaining integrity of integument, including decubitus care and massage." A fairly straightforward case could be made that, depending upon one's perspective when abstracting information, the care of a venous stasis ulcer could be correctly coded as target 01, 07, or 51. More troublesome is the fact that once the underlying nursing actions data are encoded (ie. abstracted) into a particular target intervention, there is no way to reverse the lossy data transformation that has occurred.

Georgetown Home Healthcare Classification
The nursing intervention portion of the Georgetown Home Healthcare Classification System contains 160 nursing interventions organized across 20 "care components". Each intervention is modifiable by one of four modes of intervention actions: assess, teach, provide direct care, and manage the delivery of care. Thus, the Georgetown System can rightfully claim that its vocabulary represents 640 distinct nursing interventions.

The problem with using the Georgetown system to classify nursing actions is similar to that described above with the Omaha System, i.e. that it requires nurses at the point-of-process (or worse, an after-the-fact chart reviewer) to assign atomic nursing actions to the proper Georgetown intervention classification. The fact that each of the Georgetown interventions is defined using the phrase "Actions taken to . . ." underscores the fact that a Georgetown nursing intervention is viewed as an abstraction of process, rather than process itself.   In fact, it is this ability

for abstraction that makes both the Georgetown and Omaha Systems highly useful for encoding more abstract measures such as nursing intensity, which is predictive of resource needs, but difficult to use to capture the more fine-grained aspects of nursing actions *per se*.

For example, if a nurse changes the sheets on the bed of a diaphoretic patient and administers an antipyretic, such actions could reasonably be encoded as Comfort Care - actions performed to enhance or improve well-being. Likewise, the same nurse could administer morphine sulfate and psychological counseling to a patient with terminal lung cancer and those actions could also reasonably be encoded as Comfort Care.

We do not suggest that the encoding of either set of nursing actions as Comfort Care is incorrect. Rather, it is the uni-directional (and therefore lossy) data transformation from specific atomic nursing actions to abstracted information that concerns us. Alternatively, if each set of atomic actions had been encoded using a controlled vocabulary and an appropriate compositional vocabulary designed to capture atomic nursing concepts, both informational "versions" of the underlying raw data could be viewed as "correct" in that they each represent appropriate contextual abstractions of the data.

Nursing Intervention Classification
The Nursing Intervention Classification (NIC) system from the University of Iowa has recently emerged as the leading candidate for *de facto* standard for encoding nursing interventions..   The recently expanded current (1996) version of NIC lists 433 nursing interventions partitioned into 6 domains (Level 1) and 27 classes (Level 2). For each intervention, the system provides a definition and a set (usually 10 or more) of "nursing activities" that characterize the actual intervention.

The data versus information problems surrounding NIC are somewhat more subtle than those of Georgetown because of the fact that NIC contains not only nursing interventions, but also associated "nursing activities" which are considered to be a level of abstraction more granular than interventions. If the nursing activities for each nursing intervention are considered to be "definitional" in the sense that they *are* the atomic data, then one could argue that the transformation from the set of nursing activities to the NIC code is a lossless transformation. In this

case, the NIC system would appear to at least superficially satisfy the requirements for a controlled vocabulary, albeit one with very complex atomic "words" and no associated compositional grammar.

However, a closer examination of NIC reveals substantial semantic "fuzziness" in the exact definitions of the atomic "words." In particular, nursing activities with similar semantic content can appear in more than one nursing intervention - e.g. "develop a behavior change program" appears in Behavior Modification, while "discuss with patient why the (sexual) behavior or verbalization is unacceptable; provide the predetermined consequences for undesirable (sexual) behavior" appears in Behavior Management: Sexual.

Ozbolt[8] has pointed out that NIC's reliance on "definitional" nursing activities carries the potential weakness of not capturing a key element of atomic data *logically* included by the nurse to be part of the NIC intervention, but not included on the activity list. If this problem is addressed by allowing NIC users to "customize" a NIC intervention by editing the definitional activity list for that intervention, NIC becomes, in the extreme, exactly like Georgetown, i.e. a collection of nursing intervention "bins" into which individual nursing activities are abstracted.

In summary, NIC, like Georgetown, is at core an information abstraction system rather than a controlled vocabulary on which information abstraction can be based.

Criteria for Multi-Purpose Controlled Vocabulary
Cimino[2] has identified nine criteria that collectively characterize a multi-purpose controlled vocabulary for clinical information systems. Aimed at increasing the sensitivity and specificity of information retrieval queries, the criteria are as follows: domain completeness; nonredunancy (one and only one way to express each concept); nonvagueness; nonambiguity (terms must refer to only one concept); synonymy (support multiple nonunique names for concepts); explicit relationships among terms; consistency of relationships; multiple classifications; and consistency of view.

The published literature related to the development of the three nursing classification systems recognized by the ANA demonstrates that the primary strength of these systems is that they are clinically relevant and that they meet the criteria of domain completeness for the domain of nursing for which they were specifically designed. However, it should be emphasized that vocabularies designed from the perspective of reimbursement do not, *a priori*, qualify as vocabularies capable of capturing the details of process.

The ANA Steering Committee's criteria of "precisely defined terms" with "unique codes" are explicated in more detail in three of the criteria specified by Cimino: nonredundancy, nonvagueness, and nonambiguity. In contrast to many other less rigorously developed classification systems, the three recognized systems include definitions for each of the interventions. However, the absence of a robust underlying data structure and clearly defined data transformation rules (i.e. knowledge formalisms), makes guaranteeing the criteria of nonredundancy, nonvagueness, and nonambiguity in a given data abstraction virtually impossible.

The criteria of synonymy is more appropriately applied to atomic-level data than to systems designed to cluster concepts. Data abstractions, such as those inherent in classification systems, by their nature, have difficulty with synonymy, which is much easier to control at a finer level of granularity.

The remaining three criteria (multiple classification, consistency of view, and explicit relationships) represent important areas for research aimed at refining the structure, as opposed to the domain completeness, of the existing systems. The efforts of the NLM in developing the UMLS Semantic Network and its associated vocabulary of inter-concept relationships are a major step in this direction.[10]

DISCUSSION

Graves[11] introduced the conceptual model of a three-tiered hierarchy of data - information - knowledge as a framework for conceptualizing nursing informatics. In this model, both information and knowledge represent successive abstractions and transformations of atomic data. A primary feature of any data transformation is its characterization as lossless – the abstraction does not destroy the underlying data--, or lossy – the abstraction results in data loss. Data loss can be prevented if the transformation is well-defined to be reversible. Alternatively, data loss can be prevented if the raw data is itself encoded prior to input into the transformation.

We believe that controlled vocabularies are essential in understanding nursing actions because of their ability to ground data abstractions (i.e. information) on reversible (and therefore lossless) data transformations. The present nursing intervention classification systems, while effective in presenting an abstraction of the underlying process, develop this abstraction without the benefit of an underlying atomic data infrastructure. The abstraction is thus, by its very nature, lossy. Building the abstraction on a controlled vocabulary would allow the atomic data to be recast as a different abstraction given a different set of semantic requirements. We believe that taxonomic vocabularies for nursing activities would augment and refine, rather than replace, the existing classification systems.

In this paper, we have presented evidence which we believe indicates the need for the development of a comprehensive controlled vocabulary for capturing atomic-level nursing action data. Such a vocabulary does not exist in total today. Previous work[12] has indicated that SNOMED III captures a portion of nursing process, but that it is particularly weak in areas involving nursing intervention/nursing activity. We would propose that SNOMED requires significant additions to make it "nursing responsive."

However, the existence of a vocabulary *per se* does not solve the atomic data representation problem. (Imagine a dictionary of 10,000 words and no rules for combining the words to express ideas, concepts). A compositional grammar which allows nurses to use the controlled vocabulary to express their process will be required. Examples of such grammars have been shown to be powerful adjuncts to controlled vocabularies such as SNOMED.[13]

We believe that the development of a controlled nursing vocabulary is the next major step that needs to be taken on the path to expanding nursing knowledge about clinical process. Only after such a vocabulary is developed and put into use via appropriate compositional grammars or other knowledge formalisms can information abstraction methodologies such as the nursing classification systems be used to their fullest potential.

References

1. Ingenerf J. Taxonomic vocabularies in medicine: The intention of usage determines different established structures. MedInfo 95, Vancouver, BC, 1995.
2. Cimino JJ, Hripsak G, Johnson SB, Clayton PD. Designing an introspective, multi-purpose, controlled medical vocabulary. Thirteen Annual Symposium on Computer Applications in Medical Care. 1989:513-518.
3. McCormick KA, Lang N, Zielstorff R, Milholland DK, Saba V, Jacox A. Toward standard classification schemes for nursing language: Recommendations of the American Nurses Association Steering Committee on Databases to Support Clinical Nursing Practice. JAMIA 1995; 1:421-427.
4. Martin KS, Scheet NJ. The Omaha System: Applications for community health nursing. Philadelphia: W.B. Saunders, 1992.
5. Saba VK. The classification of home health care nursing: Diagnoses and interventions. Caring 1992; 11(3):50-56.
6. Iowa Intervention Project. Nursing Interventions Classification, 2nd Edition. St. Louis: Mosby, 1996.
7. Grobe SJ. Nursing intervention lexicon and taxonomy: Preliminary categorization. MedInfo92. Geneva: North-Holland, 1992.
8. Ozbolt JG, Fruchtnicht JN, Hayden JR. Toward data standards for clinical nursing information. JAMIA 1994;1:175-185.
9. Rector AL, Nowlon WA, Kay S, Horan B, Wilson A. Foundations of an electronic medical record. Methods of Information in Medicine 1991; 30:179-186.
10. Lindberg DAB, Humphreys BL. Toward a unified medical language system. Medical Informatics Europe '89. Berlin, Germany: Springer-Verlag, 1989;1:23-31.
11. Graves JR, Corcoran S. The study of nursing informatics. Image 1989;4:227-231.
12. Henry SB, Holzemer WL, Reilly CA, Campbell KE. Terms used by nurses to describe patient problems: Can SNOMED III represent concepts in the patient record? JAMIA 1994;1:61-74.
13. Campbell KE, Musen MA. Representation of clinical data using SNOMED III and conceptual graphs. Sixteen Annual Symposium on Computer Applications in Medical Care, 1992;16:354-358.